



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2015

Simultaneous planning and action: neural-dynamic sequencing of elementary behaviors in robot navigation

Billing, E ; Lowe, R ; Sandamirskaya, Y

DOI: <https://doi.org/10.1177/1059712315601188>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-121668>

Journal Article

Accepted Version

Originally published at:

Billing, E; Lowe, R; Sandamirskaya, Y (2015). Simultaneous planning and action: neural-dynamic sequencing of elementary behaviors in robot navigation. *Adaptive Behavior*, (5):243-264.

DOI: <https://doi.org/10.1177/1059712315601188>

Simultaneous Planning and Action: Neural-dynamic Sequencing of Elementary Behaviours in Robot Navigation

Adaptive Behavior
000(00):1–13
©The Author(s) 2010
Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI:doi number
<http://mms.sagepub.com>

Erik Billing*

Interaction Lab, University of Skövde, Skövde, Sweden

Robert Lowe

Interaction Lab, University of Skövde, Skövde, Sweden

Department of Applied IT, University of Gothenburg, Sweden

Yulia Sandamirskaya

Institute of Neural Computation, Ruhr-Universität Bochum, Bochum, Germany

Abstract

A technique for *Simultaneous Planning and Action (SPA)* based on *Dynamic Field Theory (DFT)* is presented. The model builds on previous work on representation of sequential behavior as a sequence of attractors in dynamic neural fields. Here, we demonstrate how chains of competing attractors can be used to represent dynamic plans towards a goal state. The present work can be seen as an addition to a growing body of work that demonstrates the role of DFT as a bridge between low-level reactive approaches and high-level symbol processing mechanisms. The architecture is evaluated on a set of planning problems using a simulated e-puck robot, including analysis of the system's behavior in response to noise and temporary blockages of the planned route. The system makes no explicit distinction between planning and execution (in terms of neural activation). This allows continuous adaptation of the planned path. The proposed architecture exploits the DFT property of stability in relation to noise and changes in the environment. The neural dynamics are also exploited such that *stay-or-switch* action selection emerges where blockage of a planned path occurs: stay until the transient blockage is removed *versus* switch and replan to an alternative route to the goal.

Keywords

Dynamic Field Theory, Goal directed behavior, Simultaneous Planning and Action

* Corresponding author; Erik Billing, University of Skövde, SE-Box 408, 541 28 Skövde, Sweden. E-mail: erik.billing@his.se, phone: +46500-448367.

1. Introduction

The question of how an agent represents and selects actions in order to reach a behavioral goal runs deep in the fields of cognitive science and adaptive behavior, and it is an important problem for both biological and artificial cognitive systems. This problem has been approached by breaking down the behavioral repertoire of the agent into *Elementary Behaviors (EBs)* – ‘motor primitives’ or ‘schemata’ that constitute the basic action vocabulary of a cognitive controller (e.g. Matarić, 2000). This idea takes root in Arbib’s (1985) early work on distributed motor control and is central in behavior-based robotic architectures (Arkin, 1998; Brooks, 1991). While the exact formulation varies significantly between different bodies of work (e.g., Matarić, 1997; Nicolescu, 2003; Tani et al., 2004; Billing and Hellström, 2010), almost all formulations include both sensor and motor aspects of the behavior in each elementary behavioral unit. Each EB must be initiated at the right time, and there is also a need to decide when, and if, a behavior has been successfully executed. Moreover, such a distributed controller needs to coordinate competing behaviors and elaborate a sequence of behaviors that leads to a long-term behavioral goal.

One way to approach this problem is to formulate EBs as attractor patterns in a dynamical system that may be coupled to the agent’s sensorimotor dynamics. Such an approach has the advantage that planning decisions of selecting, activating, and terminating EBs and sensorimotor dynamics of the individual behaviors are formulated in the same computational substrate of attractor dynamics, which leads to a more homogeneous architecture. Our work builds on one such formulation (Sandamirskaya and Schöner, 2010; Sandamirskaya et al., 2011) within the framework of Dynamic Field Theory (DFT). In earlier work, we considered the *serial ordering* of actions in time (Sandamirskaya and Schöner, 2010) and the flexibility in the order of actions *behavior organization* (Richter et al., 2012), introducing EBs within the DNF framework (Sandamirskaya et al., 2011). Here, we extend the architectures for sequence generation to a distributed dynamical controller capable of *planning* and executing a sequence of EBs that leads to a behavioral goal. To study the functionality of the new architecture, we implement a path planning scenario for a simulated robotic vehicle, as an example of a more general planning and search capability. Using the extended DNF-based EB formulation, we enable path search and planning towards a given goal, robust acting out of the planned trajectory, and updating of plans if the environmental situation changes. To our knowledge, this work represents the first attempt to utilize the neural-dynamic DFT framework for path planning and search.

We will present the basic concepts of the DFT framework, required to understand our work and our motivation for using this particular formalism, in the next section. The relation of our work to other planning and search algorithms will be considered in the Discussion.

The focus of the present work concerns linking the lower level sensorimotor dynamics to higher cognitive processes as they manifest themselves in planning and search. This is achieved through a biased competition between the currently viable action candidates and, simultaneously, between the future action candidates. With this approach, the local sensory-driven (reactive) behaviors are integrated with the ‘long-term’ plan, represented as an attractor pattern over a number of Dynamic Neural Fields. The approach we present here enables us to deal with the problem of shallow gradients, which is otherwise common in, e.g., potential field approaches. In the proposed architecture, attractors are sustained during action, both in the sensory-motor target DNFs and at the planning level, allowing for robust performance within a changing environment. This feature also permits long trajectories to remain stable in the face of transient sensory perturbations. Furthermore, our DNF approach has the potential to maintain many plans and sub-plans in parallel whilst simultaneously carrying out actions. To this end, we refer to our approach as implementing *Simultaneous Planning and Action (SPA)*. This absence of separation of planning (as well as path extraction) and execution phases, we suggest, promotes great flexibility and robustness of performance of the robotic agent whilst not severely compromising on measures of optimality (e.g. shortest path finding and maintaining safe distances from obstacles).

To the authors’ knowledge, no DNF-based system has been proposed that searches for a sequence of actions towards a goal, in a way that can be compared to classical AI formulations of planning (as we do in the Discussion). While efficient and optimal search may have been overemphasized by classical AI in explaining human and animal cognition, search is without question a key component in at least higher level cognitive functions, such as planning. For this reason, we believe

that a DNF formulation of search and planning can contribute to a better understanding of the link between the lower-level sensorimotor processes and higher-level cognitive functions. It may also shed light on the pros and cons of the dynamical, sensorimotor, or embodied approaches in relation to classical AI-formulations of search.

The rest of the paper breaks down as follows. After introducing the basics of DFT in Section 2, we present the DNF architecture able to plan a sequence of actions leading towards a goal in Section 3. The presented architecture should be understood as a generalization of previous work on sequence representation (Sandamirskaya et al., 2011; Sandamirskaya and Schöner, 2010). Three test cases applying the architecture as a planning and control system for a simulated e-puck robot are presented in Section 4. These illustrative demonstrations were selected to give an intuitive understanding of the architecture’s behaviour. While the presented work should be understood as a generic search architecture, we make an effort to present a complete system, where inputs and outputs interface directly with the world. Results from simulated robot experiments are given in Section 5. We here focus solely on planning within an existing state space rather than on problems of localization and mapping (i.e. construction of a cognitive map of the environment). A discussion of possible connections to animal planning and learning abilities, as well as of the relation of our work to AI planning and search algorithms are presented in Section 7.

2. Methodological background: Dynamic Neural Fields

DFT (Schöner, 2008) is today a well established neurally-based framework in cognitive science, used to model various perceptual (Johnson et al., 2008; Zibner et al., 2011), motor (Schöner et al., 1997; Bastian et al., 2003), and cognitive functions (Spencer and Schöner, 2003; Schöner, 2008; Sandamirskaya et al., 2013). DFT is often presented as a bridge between the sensorimotor levels of neural processing and levels that relate to cognitive processes (Spencer et al., 2009).

In the language of DFT, the state of a cognitive system is characterized by dynamic activation functions – Dynamic Neural Fields (DNFs). A DNF is a mathematical formulation of the neuronal activation at population level, taking root in the mean-field approximation of the activation dynamics in biological neuronal networks (Wilson and Cowan, 1973; Amari, 1977; Ermentrout, 1998). DNFs are defined over behaviorally relevant dimensions, for instance, continuous perceptual features (color, orientation, or location) or motor parameters (joint position, velocity, or force), or discrete cognitive dimensions (serial order, object labels). The activation of a DNF evolves in time according to Equation (1), first analysed by Amari (1977). This dynamical system equation has an attractor solution of a particular shape – a localized bump of positive (suprathreshold) activation. This bump, or peak, is stabilised by the recurrent interactions in the DNF. The position of the suprathreshold activation bump on the DNF’s dimension specifies the content of the respective representation, i.e., which parameter values characterise the current state of the cognitive system. The strength of the activation bump expresses the certainty of the system in the current estimation of the behavioral parameter. Such localized activity bump, or peaks, are units of representation in cognitive architectures, built with DNFs (Schöner, 2008; Sandamirskaya et al., 2013): they represent perceptual objects, motor intentions, or plans.

The temporal dynamics of the activation function of a DNF, $u(x, t)$, defined over a continuous behavioural dimension, x , is formalised in Eq. (1):

$$\tau \dot{u}(x, t) = -u(x, t) + h + \int f(u(x', t)) \omega(x - x') dx' + S(x, t), \quad (1)$$

where τ is the time constant of the field dynamics, t is time, $h < 0$ is the resting level that ensures that without external input the field is subthreshold, i.e. at a negative activation level, and $S(x, t)$ represents the external stimuli. The connectivity function $\omega(x - x')$ is a bell shaped kernel representing a short-range excitatory (c_{exc}), a longer-range inhibitory (c_{inh}), and a globally inhibitory (c_{global}) inflows from other activated locations in the field:

$$\omega(x - x') = c_{exc} \exp \left[-\frac{(x - x')^2}{2\sigma_{exc}^2} \right] - c_{inh} \exp \left[-\frac{(x - x')^2}{2\sigma_{inh}^2} \right] - c_{global}, \quad (2)$$

where σ s are widths of the respective parts of the kernel.

The output of the DNF is shaped by a sigmoidal non-linearity, f , which determines, which locations of the DNF provide input to other locations in the field, and possibly other connected fields:

$$f(u(x, t)) = \frac{1}{1 + \exp[-\beta u(x, t)]}, \quad (3)$$

where β is the slope of the sigmoidal non-linearity.

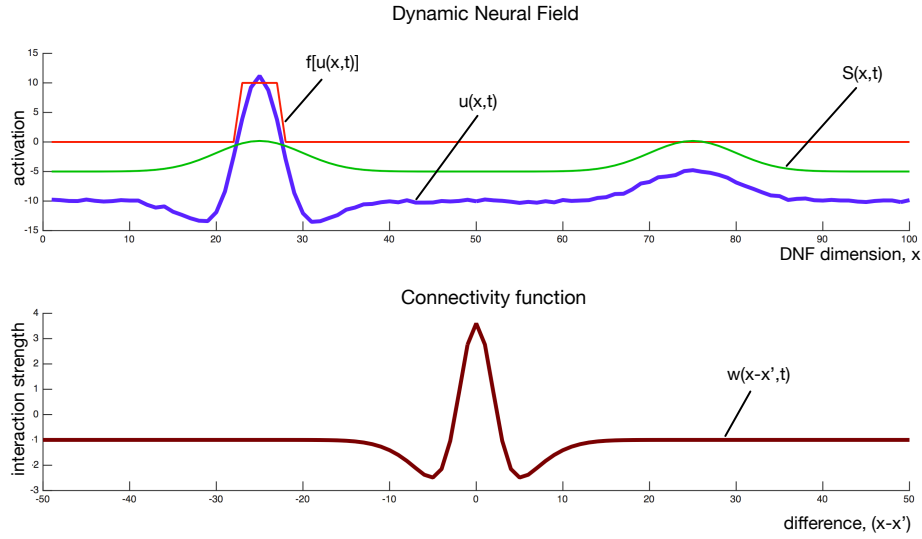


Figure 1. An exemplary one-dimensional DNF (Eq. 1; top) and the respective interaction kernel (Eq. 2; bottom). The activation level $u(x, t)$ – blue line, sigmoid output $f(u(x, t))$ – red line, and external input to the field $S(x, t)$ – green line, are shown.

Fig. 1 shows a one-dimensional DNF, described by Eq. (1), and the respective lateral interactions kernel, Eq. (2). The external input, $S(x, t)$, to the field (green line in the plot) has two regions of higher strength, in which the activation function of the DNF (blue line in the plot) reaches the activation threshold (zero line). The lateral interactions of the DNF, described by the kernel function depicted in the lower part of the figure, “pull” a localized activity peak over one of the regions with high input strength (the one that reached the activation threshold earlier), inhibiting the DNF at other locations. This inhibition also suppresses activation in the second region with high input strength.

Thus, through its dynamics, the DNF here made a *decision*, selecting one of the two regions in the input distribution and stabilizing this decision by lateral interactions. This decision will have impact on down-stream structures in a DNF architecture through coupling of this field to other DNFs and, ultimately, to the sensory-motor dynamics.

A different parametrisation of the interaction kernel, with $c_{global} = 0$, would lead to a DNF that builds two localised activity peaks if stimulated with the same input as the one, depicted in Fig. 1. If the excitatory part of the interaction kernel is strong enough, on the other hand, the activity peak may remain supra threshold even when the initial input to the field has vanished. Cognitive neural-dynamic architectures can be built with DNFs with different parametrisations and have been used both to model human cognitive behavior (Johnson et al., 2009) and to control cognitive robots (Sandamirskaya et al., 2013).

3. The neural-dynamic SPA architecture

3.1. The general framework

In previous work, Sandamirskaya and Schöner (2010); Sandamirskaya et al. (2011) have used the notion of *Elementary Behaviour (EB)* to describe a discrete functional component representing some part of the complete behavioural repertoire of the agent: e.g., “look for an object with color X” is one possible EB, which implements search and approach (or gaze) towards an object of a given color, “move arm to position Y” is another typical EB. Each EB is characterized by two neural-dynamic structures, related to control of action initiation and termination: *intention (I)* and *condition of satisfaction (C)*, each represented by a DNF (Sandamirskaya et al., 2011). When the I-field is activated, it impacts the downstream dynamical structures, connected to the sensors and actuators of the robot, and the intended behaviour is executed. Moreover, it pre-activates (or pre-shapes) the C-field to be sensitive to the sensory outcome of the behavior. When the sensory conditions are perceived that correspond to a successful accomplishment of the current action, the C-field becomes active, signalling that the behavioural goal is achieved. The C-field also inhibits the I-field and deactivates the EB.

Sandamirskaya et al. (2011) demonstrated how this structure enabled autonomous organization of simple action sequences. To encode order of actions in a sequence, nodes representing rules of behavioural organization were introduced. In particular, the *precondition (P)* node inhibits the initiation of an EB that requires certain conditions to be fulfilled in order to be activated. The P node is inhibited, in its turn, by an activated condition of satisfaction (C) node of a different EB. Sandamirskaya et al. (2011) have demonstrated how different action sequences may be encoded in this architecture and may be activated by selecting one of the task nodes. A task node boosts all EBs and precondition nodes, involved in achieving the selected behavioral goal. As a result, a sequence of behaviors unfolds, leading to accomplishment of a task, such as grasping, pointing, or lifting an object.

Here, we present a planning architecture that consists of four main parts, depicted in Fig. 2. Analogous to previous work, there are *intention (I)*, *condition of satisfaction (C)*, and *precondition (P)*, components, implemented as DNFs with a fourth component, *motivation (M)*, added. M should be seen as a goal representation, feeding activity to both P (connection 1, Fig. 2) and I (connection 3). I indicates the initiation of the action, but also parametrization of the sensorimotor system during the particular action (e.g. color of interest or location of the motor goal). There is consequently a close link between the goal representation (M) and the initiation of an action to reach that goal (I). Supra threshold activity in the precondition field (P) suppresses activation of I (connection 4). The precondition can be released by an inhibitory input from the perceptual system or internally, by a C node of another behavior, which signals its successful execution (connections 7).

Planning emerges in interaction between the M and P fields, where a plan is represented as a pattern of self-stable supra-threshold attractors (cf. Schöner, 2008) distributed over these two fields (connections 1 and 2). The active regions in the P fields inhibit respective regions in the I-fields, which would otherwise be activated by input from the active M-fields. When one of the P-fields receives inhibitory input (connection 7) strong enough to inhibit the relevant location in this field and thus release the inhibition of the intention (connection 4), a peak rises in one of the intention fields, and the plan unfolds into action.

An example of a minimal (trivial) planning sequence using recurrent activation of a single EB. The spreading of activation is illustrated in Figures 3 and 4. The minimal *plan* helps illustrate the spreading activation property (transition between EBs) of our DFT framework. Slices of the two-dimensional fields are depicted, illustrating the formation of a *plan* along a continuous dimension of a single EB. Each location on this dimension corresponds to action of the EB, with a different parametrization (e.g., movement to a particular location in space). Active locations in the M-field (motivation) excite the respective locations in the I- (intention) and the P- (precondition) fields. If activity in the P-field is not inhibited, it suppresses the respective location in the I-field, preventing initiation of the action.

An active part of the P-field may also activate the M-field, that is, the architecture is forming a motivation to fulfill that particular precondition. A single site of P may activate several locations in the M-field, or even M-fields of other EBs, as

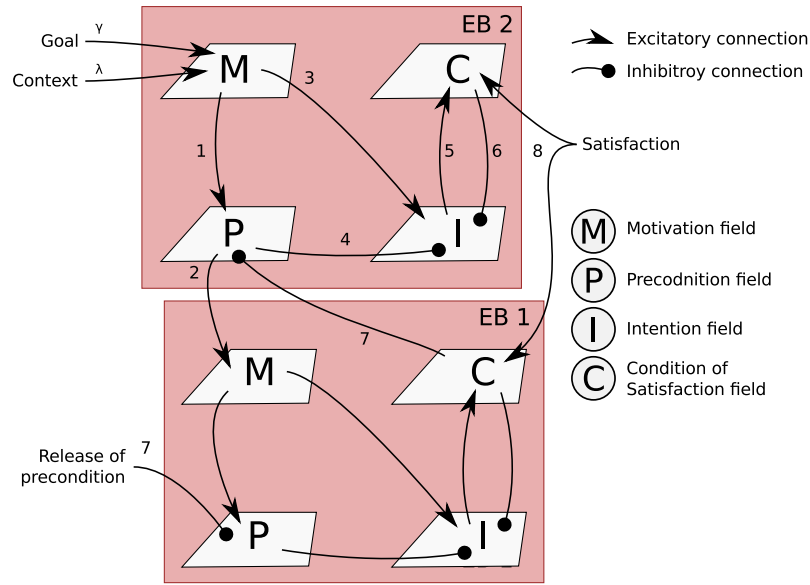


Figure 2. Schematic of the main components of the proposed planning architecture. A sequence of two elementary behaviors (EB1 and EB2) is displayed. A goal state is here represented as a motivation (m) to execute EB2, given that some preconditions (p) are fulfilled.

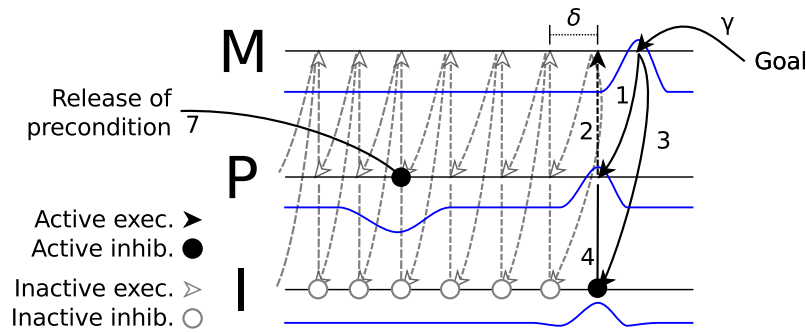


Figure 3. Initiation of a minimal plan. The plan consists of recurrent activation of a single EB, illustrated as an intersection of the 2-dimensional motivation (M), precondition (P), and intention (I) fields. An external (goal) stimulus (γ) activates a particular site at the M-field. The blue lines represent activity in the corresponding fields shortly after the goal stimulus is introduced. The black line corresponds to the detection threshold (c.f. Section 2). Detection concerns transitions from EB to EB (in this case the transition is to the same EB). Activity spreads, via connections 1 and 3, to P and I. As a peak forms in P, the I-field is inhibited (connection 4), preventing execution of the behavior. Activity will continue to spread via connection 2, eventually forming a plan between M and P. Field activity corresponding to a complete plan is illustrated in Fig. 4. δ denotes the size of the shift for connections 1 and 3.

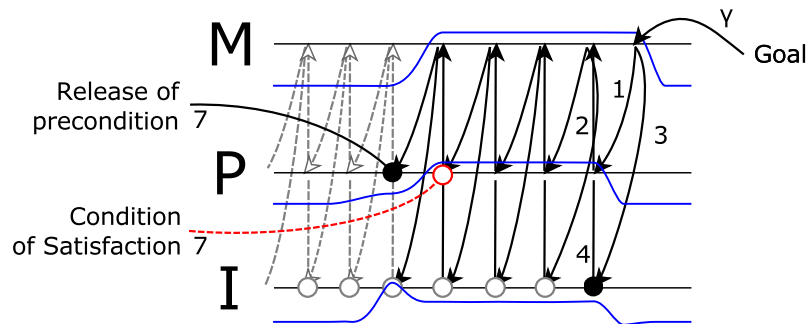


Figure 4. Completion of minimal plan. Recurrent activation of motivation (M) and precondition (P) of a single EB (compare with Fig. 3). Activity spreads until intersecting with the release of precondition, suppressing activity in P, allowing a peak to form in the intention field (I), initiating execution of the behavior. If execution is successful, a condition of satisfaction peak will form (indicated as a red input, compare with Figure 2), releasing precondition for the next step of the plan.

illustrated in Fig. 2 and later discussed in Section 3.2. Eventually the spreading motivation may lead to activation of an M-field for which the precondition is suppressed (Fig. 4), allowing a peak to form in the I-field, and the plan unfolds into action.

3.2. The SPA architecture implemented for the path planning scenario

Planning. The SPA architecture for the navigation example, considered in this work, is shown in Fig. 5. The implementation of the SPA framework for this scenario requires some explanation: the individual EBs here correspond to elementary movements in one of the directions in space (arbitrarily denoted as North, South, West, and East). Each EB is parameterized along two dimensions, corresponding to physical space, resulting in two-dimensional DNFs implementing M, P, I, and C (compare with Fig. 2). Locations in M, P, I, and C fields represent transitions between places in the environment. For instance, activity at a location in the I- (intention) field of the *North* EB initiates a movement towards the north, potentially involving the robot turning in this direction. The robot will continue to move north as long as there is activity in the *North* I-field, but the precondition for moving north will normally never be suppressed unless the robot is standing at a location from which it can move north towards a motivated goal.

Fig. 5 shows how the M, P, I, and C DNFs are connected with sensors and actuators of the robot. The I-fields propagate activity to four action fields, which induce rotations of the robot's wheels. A peak in the *condition of satisfaction* C-fields forms when the robot reaches a position adjacent to the current location, i.e., when input from connections 5 and 8 intersect. Mathematical formulations of this network are given in Appendix A. Note that the SPA framework is not limited to four EBs representing compass directions; this particular configuration was used to constitute an intuitive example comparable with classical planning approaches.

Some of our EBs are mutually exclusive (e.g., *South* and *North*). Competition between opposing EBs is implemented in a similar way as the motivational links (connection 2, Figures 2 — 5), but is represented by the inhibitory rather than excitatory connections. For example, activation in the *East* M-field representing motivation to go east towards a specific position X results in activated preconditions for locations west of X (compare with Figure 3 and 4). Activity in the *East* P-field will spread, via excitatory connections, to corresponding locations in M-fields of *North* and *South*, but also will suppress activity in the *West* M-field. This results in a cascade of competing activation between M- and P-fields, among the available set of EBs. One example of this cascading activity among all four EBs is presented in the result section (Figure 13).

The M-fields receive a *Goal* (connection γ) input – essentially a place-holder for the neural activity correlated with a reinforcing input, e.g. an anticipated or actual rewarded state. In other work (Gaussier et al., 2000; Hirel et al., 2013) motivations are considered in relation to drives (Tolman, 1948) and gate the valuation of the particular goal location. We take a similar approach by representing the goal as a stimulation to one or more M-fields. This stimulation constitutes the starting point for the formation of a plan. In the present work, a single Gaussian goal input is given to all M-fields, corresponding to a specific target location in the local surroundings that potentially can be reached from any direction.

Obstacles and other behavioral restrictions imposed by the environment are represented as inhibitory contextual input to the motivation fields (connection λ , Figure 5). The context should here be understood as a memory of the local surroundings. It should be noted that both the DNF representations, and their mapping to the world through sensors, are continuous in both space and time. A contextual or goal input may change at any time, sometimes resulting in dramatic changes in the activation patterns of the fields, and as a result, the robot's behavior. While plans form as competing activity in the M and P fields and is in this sense continuous, each EB implement a certain connection shift (δ , c.f. Fig. 3) and has in this sense a discrete component. Plans may also involve several different EBs, e.g. execute *West* to fulfill the precondition for going *North*. In an different scenario, this may involve moving towards a table before reaching for an object, where

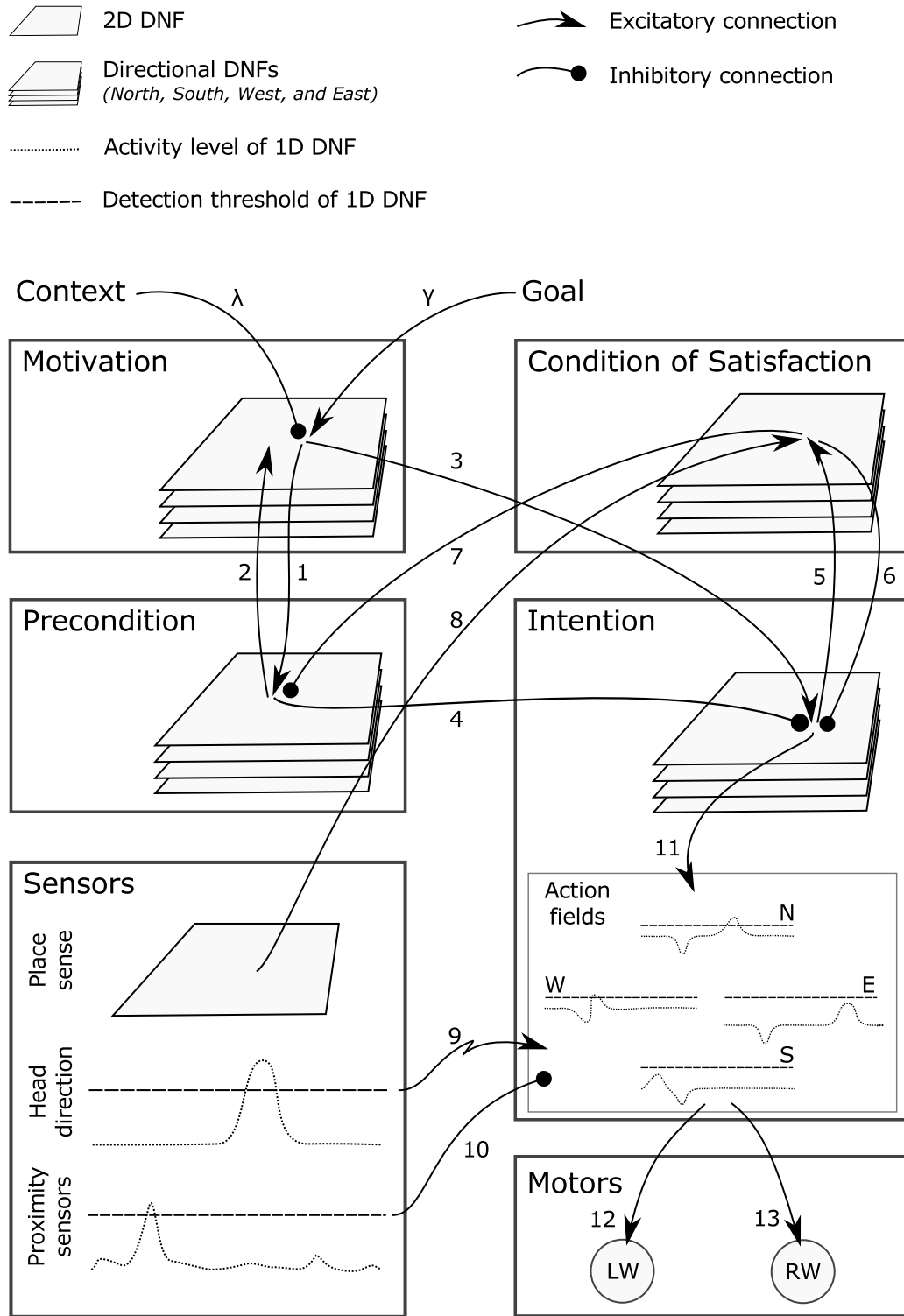


Figure 5. Overview of the complete architecture, including interactions with the robot. See text for details.

reaching and moving are formulated as different EBs. In this way, the architecture opens up for planning between different representational spaces, compare with e.g. LaValle (2006).

The agent may have many memories of different surroundings, but only one is active and imposed onto the M-fields at a single time. How these context representations may be formed and selected is discussed in Section 7, implementation of these processes is outside the scope of this work.

Sensing and Action. The SPA architecture for path planning integrates sensor information of three types, *place sense*, *head direction*, and *proximity sensors*, c.f. Fig. 5. The place sense provides sensory evidence for condition of satisfaction. In a general sense, it provides information that one EB has been successfully executed, the desired location has been reached, and the next step of the plan can be initiated. In the present implementation, the place sense field receives Gaussian input from a single noisy position sensor, see Section 4 for details.

Activity in the head direction and proximity sensor fields bias action selection (Fig. 5. Each EB comprise one *action field*, a one-dimensional DNF defined over an angular dimension, representing the turning angle of the robot. The head direction field provide a single Gaussian input to each action field (connection 9).

By default, the head direction input results in sub-threshold activity of the action field and hence no motor output. However, when the action field gets additional input from the I-field (connection 11), the combined input results in a supra-threshold peak in the action field, and motor output. The detailed mathematical formulation is given in Appendix A.

Each action field also gets inhibitory input from the proximity sensors (connection 10). This input may shift or completely suppress a peak in the action field, and as a result, adjust or prevent a turn in a specific direction. This mechanism can be seen as a neural-dynamic implementation of obstacle avoidance.

4. Robotic demonstration

The proposed architecture ¹ was implemented using the Matlab framework COSIVINA (Schneegans, 2015) and evaluated in two simulated environments, using the robot simulation software Webots (Cyberbotics, 2015). The first environment is a minimalistic Z-maze, depicted in Figure 6. The second environment, depicted in Figure 7, is a large maze introduced to test the architecture in a more complex setting. It should be noted that, despite their block-like look, these environments are continuous.

Case 1 100 pairs of start-goal locations in the Z-maze were selected randomly

Case 2 100 pairs of start-goal locations in the large maze were selected randomly

Case 3 This test case was designed to evaluate the behaviour of the system when obstacles are introduced, with fixed start and goal locations as displayed in Figure 6. Two types of obstacles were used, a partly obstructing obstacle and a fully obstructing obstacle. The test case was executed 240 times for each obstacle type.

Start and stop locations were selected with a min distance of 40 mm from walls and obstacles. The obstacle introduced in Case 3 appeared at a random time t_i and was removed at time $t_i + t_p$. t_i and t_p were drawn from a uniform distribution such that $30 < t_i < 700$ and $2 < t_p < 500$ simulation steps. These limits were selected to make the obstacle appear during planning or execution, before the robot reached the goal. Apart from introducing an object in the simulator, the map was updated with the new obstacle information, providing a new contextual inflow to the motivation fields. While the information about the obstacle is directly presented to the agent as an inhibitory input to the M field, in a way that does not

¹ All source code required to run the presented experiments is available for download at <https://bitbucket.org/interactionlab/spa>. Recorded data is available as a specific software branch <https://bitbucket.org/interactionlab/spa/branch/data>.

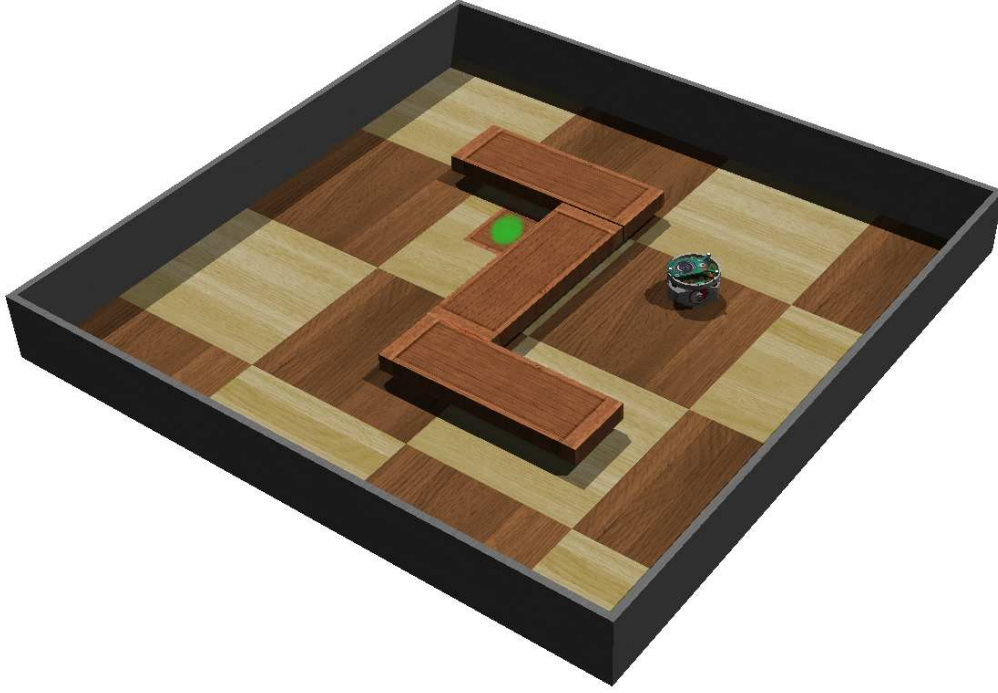


Figure 6. Evaluation environment 1, the Z-maze. The E-Puck robot is here standing at the starting point used for test case 3. The green area marks the goal.

directly correspond to any embodied perception of the obstacle, we believe it serves well to evaluate the effect of changing world information during planning and execution.

Each test case was repeated with two different levels of noise, while keeping all other parameters constant. With the low noise level, normally distributed noise with an $\text{std}=5\%$ of the contextual stimuli strength (λ in Figure 5) was applied to the map. Normally distributed noise with std of 5 mm was also applied to the position sensor. With the high noise level, std of 100% of λ and 20 mm was applied to the map and position sensor, respectively. Noise mean=0 in all cases. Distance measures in the simulated environment are given in proportion to the physical e-puck robot (Mondada et al., 2009), with a diameter of 74 mm. Locations for the obstacles introduced in test case 3 are depicted in Figure 14.

In addition to the tests using the proposed architecture, test cases 1 and 2 were executed with a reference implementation of *Follow the carrot* (e.g. Barton, 2001). For details on the implementation of *Follow the carrot*, please refer to Appendix B.

5. Results

5.1. Results for test case 1 and 2

The proposed SPA architecture was compared to a standard path tracking approach *Follow the carrot*. Tests were executed in two environments, a minimalistic z-maze (Case 1, Fig. 6) and a large environment (Case 2, Fig. 7). Test cases 1 and 2 comprised a total of five conditions each, the proposed SPA architecture with low and high noise, and *Follow the carrot* tested without noise, with low noise, and with high noise. Each condition was executed 100 times for at most 3000 simulation steps (corresponding to 192s with a simulated time step of 64 ms). This limit was selected to be well above the time necessary to reach the goal, even for the longest paths.

The proportion of successful runs is presented in Fig. 8. To get an overview of the stability and optimality of the proposed architecture, divergence from the optimal path was calculated as L_p/L_o . L_p represents the executed path length

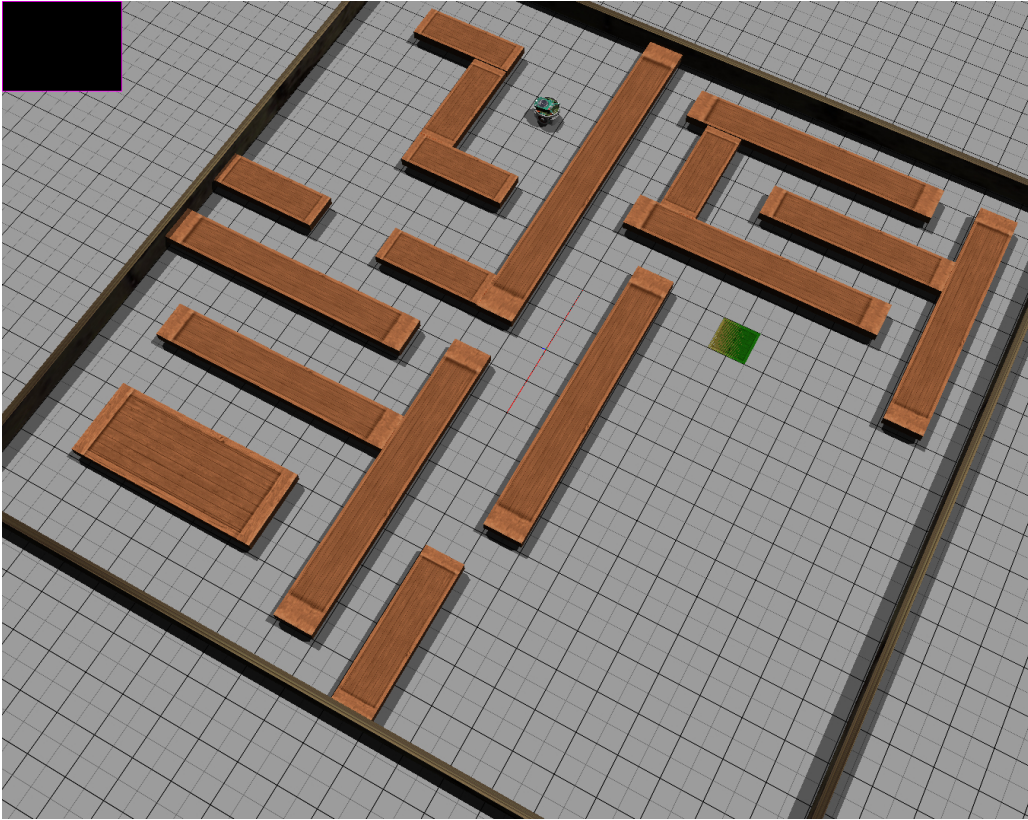


Figure 7. Evaluation environment 2, the large maze used for test case 2.

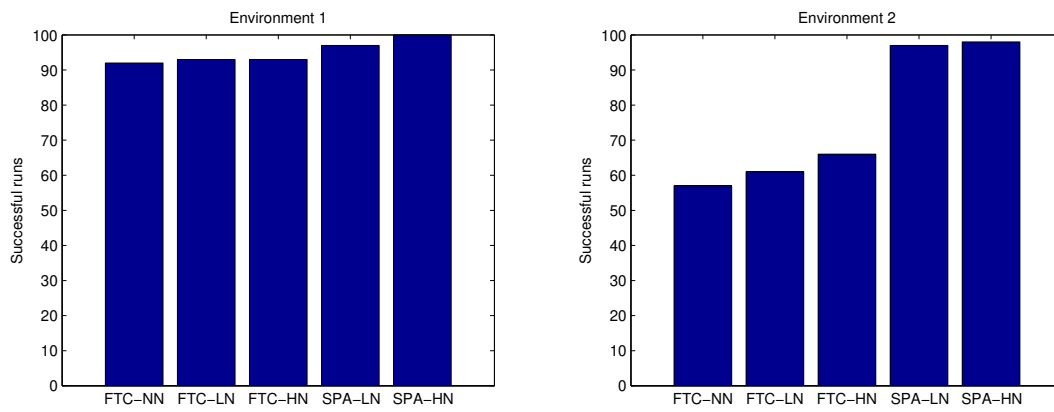


Figure 8. Number of successful runs for environment 1 (left) and 2 (right). Results are given for the proposed SPA architecture tested with low noise (SPA-LN) and high noise (SPA-HN), compared to *Follow the carrot* without noise (FTC-NN), with low noise (FTC-LN), and high noise (FTC-HN).

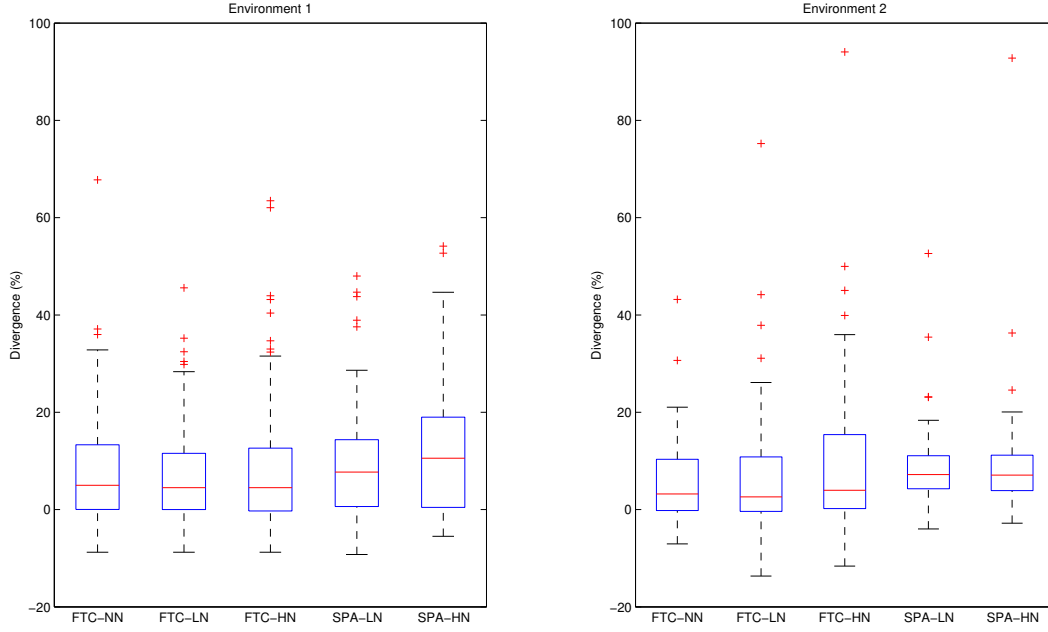


Figure 9. Divergence from optimal path for environment 1 (left) and 2 (right). Results are given for the proposed SPA architecture tested with low noise (SPA-LN) and high noise (SPA-HN), compared to *Follow the carrot* without noise (FTC-NN), with low noise (FTC-LN), and high noise (FTC-HN).

and L_o is the optimal path, defined as the Euclidean distance given by A* over an 8 neighbor grid with a cell size of 1 cm and a minimum obstacle distance of 4 cm. A One-way ANOVA revealed no significant differences in path length between different conditions in environment 1 ($F=0.97$) and environment 2 ($F=0.98$).

Executed paths produced by SPA from six representative runs of Case 2 are presented in Figure 10.

5.2. Results for test case 3

In order to analyze the detailed function of the architecture, activity in all fields was logged during a single run from Case 3, where a fully obstructing obstacle was introduced at $t = 336$. Figure 11 displays the temporal evolution of activity at four locations in the motivation field, which are defined below, for each of the four EBs (*North*, *East*, *South*, and *West*). The same four locations are displayed as a path plot in Figure 12. A snapshot of the activity in the motivation fields at the moment when planning is complete and the robot starts to move ($t = 210$) is presented in Figure 13.

The robot starts at *Location 1* ($t = 0$), then continues to *Location 2* ($t = 500$), *Location 3* ($t = 850$) and finally *Location 4* ($t = 1250$). At $t = 0$, an excitatory input is introduced in all motivation fields, at the location of the goal. Through the interplay between the M and the P fields, motivation is back-chained for each action. I.e., activity spreads north in the south motivation field, and vice versa, but also from the south motivation field to the east and west fields, leading to a competitive spread of activation in all directions. As a consequence, the goal input first propagates to the north and east motivation fields at Location 4, see in Figures 11 and 12. The south and west M-fields receive inhibitory input from north and east, respectively, leading to the dip in activity at this location. Around $t = 150$, the wave of propagating activity reaches Location 2, activating the west motivation field, and at $t = 210$, the complete “plan” is present as a pattern of suprathreshold activation in the four motivation fields (Figure 13).

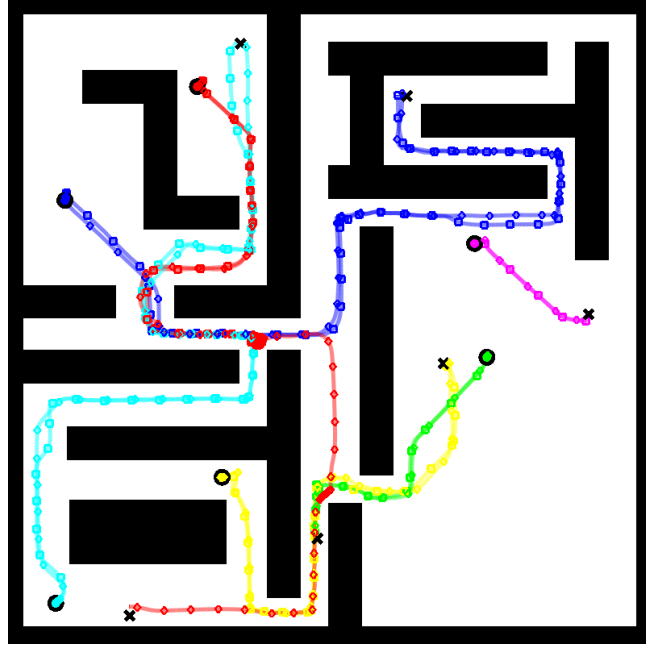


Figure 10. Path plot of six representative runs from Case 2, each start/stop pair plotted with a unique colour. Lines are marked with diamonds and squares, representing low and high noise, respectively.

After $t = 210$, activity increases in the intention fields (Figure 11, bottom pane). As a consequence of strong activity in the north field, and competing activity in the west and east fields, the robot leaves location 1, turns north, and as the west field wins the competition, continues towards north west (Figure 12).

A fully obstructing obstacle is introduced during early execution ($t = 336$). The approximate interval $350 < t < 550$ can be seen as a replanning phase where the activity in the west motivation field dies away, and is replaced by activation in the south and east motivation fields, at Location 2. The resulting sequence of actions is visible in the maximum activity of the four intention fields (Figure 11, bottom pane). Around $t = 550$, the east attractor stabilizes at Location 2, leading high activity in the east intention field, followed by south, west, and finally north.

This replanning behavior is dependent on the parametrization of the neural fields, resulting in self stable, but not self sustained, field activation. That is, active regions of the field that no longer gets propagated support from a goal will decay. This is the case when the obstacle is introduced, cutting propagation of motivation to go west, and as a result, activity at location 2 of the west M-field (figures 11 and 12) will decay, allowing competing motivation from the south to take over, resulting in the initiation of an alternate route.

An overview of the systems response to introduced obstacles is presented in Figure 14. Individual runs are colored according to the selected path, where green represents a direct route north of Z (shortest), blue represents a direct route south of Z (when the north path is blocked), and red represents an initial selection of the north path with a change to the south.

The time course for some representative runs from Case 3 is presented in Figure 15. If no obstacle is present, the robot would always go for the shortest (green) path, not displayed here. When the obstacle appears during early planning, the robot goes instead for the slightly longer south (blue) route, independently of obstacle type (Figure 15, top pane). Obstacles appearing later do not have the same effect. A partly obstructing obstacle appearing during late planning or execution will not result in a change of the selected path. The robot persists on the north (green) path, which should be understood as a case of path adjustment rather than complete replanning. However, when a fully obstructing obstacle appears during late planning or early execution, the robot switches from the north to the south route (red). A period of velocities close to zero can in these cases be understood as the time of replanning.

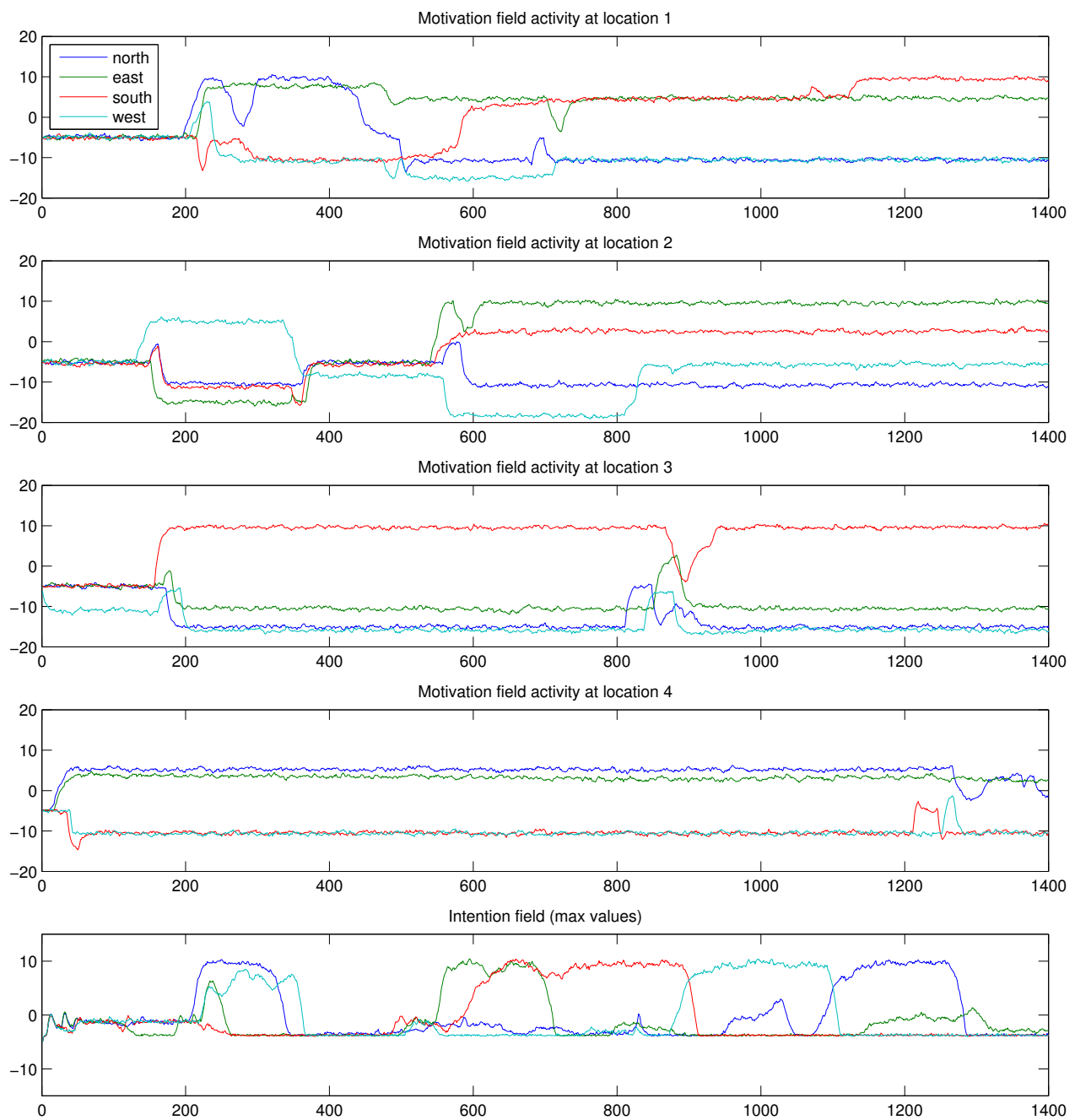


Figure 11. Field plots from the run displayed in Figure 12. The top four plots present activity from motivation fields, at four locations the robot visited during execution. The bottom plot presents maximum activity in intention fields.

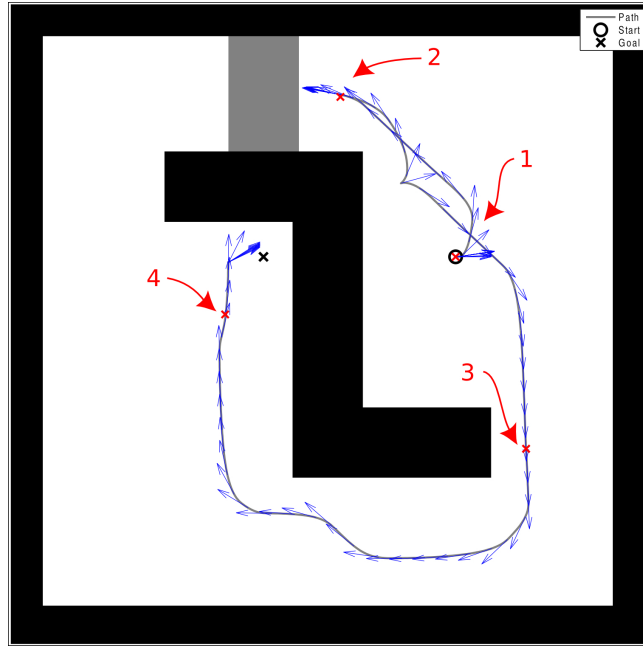


Figure 12. Path plot over a single run from test case 3. Red x marks positions that the robot passed during execution. The black x marks the goal. The grey area represents the location of a fully obstructing obstacle introduced at $t = 336$.

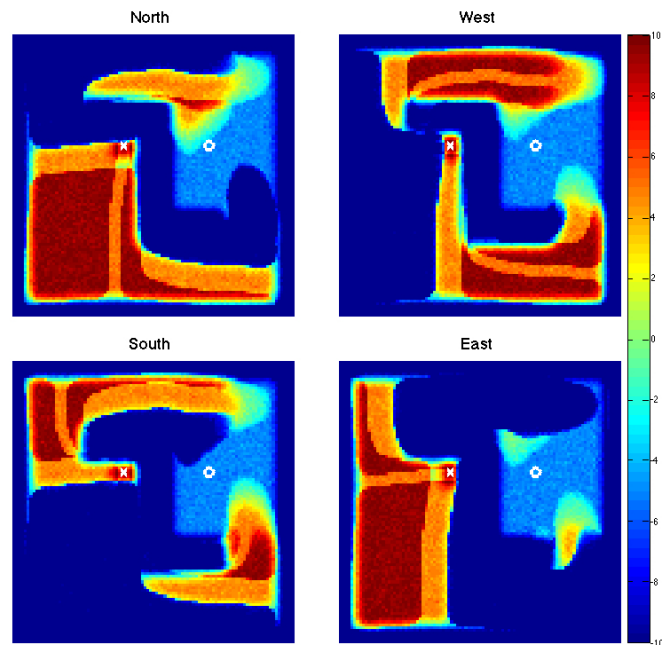


Figure 13. Activity plot over motivation fields at $t = 210$. The pattern of supra-threshold activity (red) can be seen as the plan of moving north, west, south, and finally east from the starting point (white circle) to the goal (white cross).

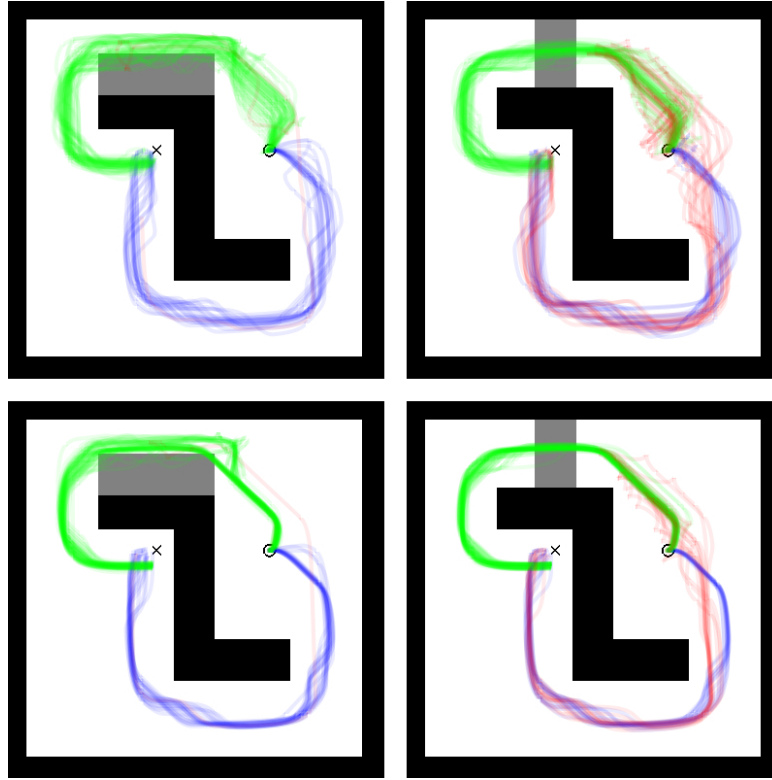


Figure 14. Path plots over runs from Case 3, where an obstacle was introduced at a random time. Left and right plots represent partly and fully obstructing obstacles, respectively. Upper plots show runs with high noise, and lower display runs with low noise. Colours represent path selections: green represents a direct route north of Z (shortest), blue represents a direct route south of Z (when the north path is blocked), and red represents an initial selection of the north path with a change to the south.

A detailed analysis of path selection and replanning behavior is presented in Figure 16 and 17. As visible in Figure 16, the appearance of a partly obstructing obstacle during the first 200 time steps results in the selection of the south (blue) route, even if the duration of that obstacle is short. However, the fully obstructing obstacle that persists for a short time (Figure 17) appears to have less effect on the decision making (i.e., the robot selects the north route to a larger degree).

When the obstacle occurs after $t = 200$ time steps, which roughly corresponds to the time the robot starts moving (see Figure 15 for details), the robot makes an initial selection of the shortest, north, path in almost all cases. If a fully obstructing obstacle appears within $200 < t < 500$, the robot switches from the north to the south route (red). The robot is, however, more persistent in continuing on the north (green) route compared to runs where the obstacle appeared before $t = 200$, that is, the robot waits longer for the obstacle to disappear if the obstacle appears after execution started. For runs where the obstacle appears after $t = 500$, the robot always selects the north route. In these cases, the robot has already passed the location of the obstacle at the time it appears, and it therefore does not interfere with the executed plan.

6. Discussion

6.1. Related work

Commonly, ‘global’ planner algorithms revolve around classical graph- and tree-based search algorithms (LaValle, 2006; Russell and Norvig, 1995). Typically, graph-based approaches entail a *search* and an *execution* component. The particular application requirements constrain the choice of the approach; for example, where robust performance (avoiding collisions and maintaining accurate path integration) is concerned, constructing a Voronoi diagram that reduces likelihood of obstacle interference may be desirable. Alternatively, if the emphasis for performance is on finding optimal paths to targets, dividing

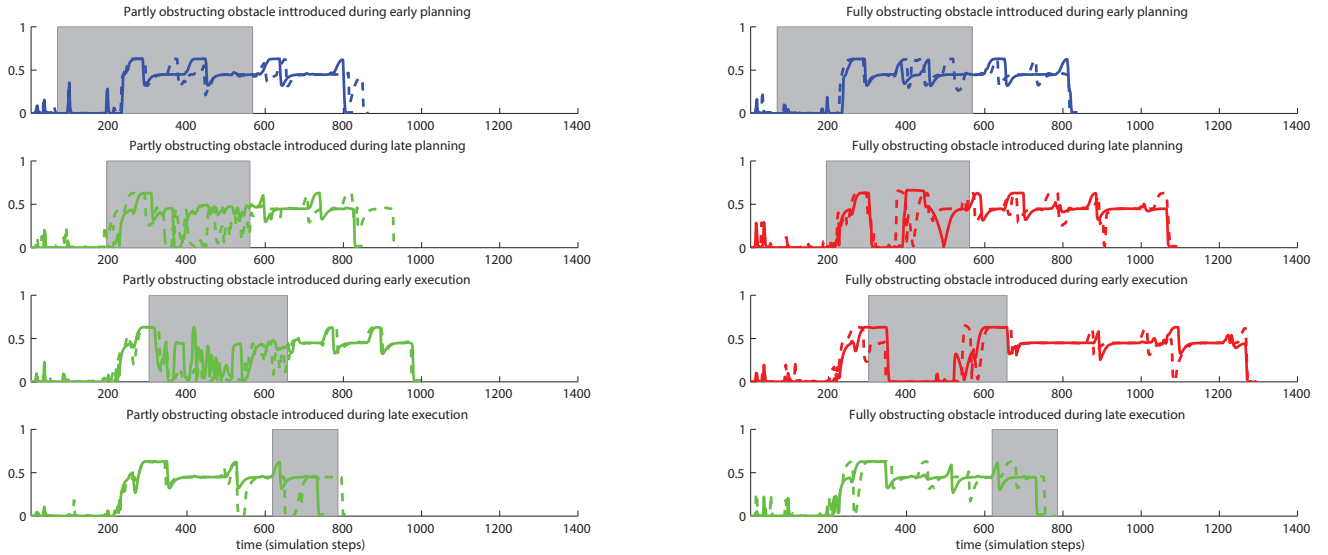


Figure 15. Velocity plots from four representative conditions in Case 3, where an obstacle is introduced during early planning, late planning, early execution, and late execution, in order from top to bottom. The grey area represents the time that the obstacle was present. Solid and dashed lines represent the velocity over time, with low and high noise, respectively. Lines are coloured according to selected path, same as Figure 14.

the state space appropriately for computing the shortest possible path may be desirable, e.g. Distance Transform (DT) methods. In the latter case, algorithms exist that are capable of finding viable and optimal paths at low computational cost. They are used for ‘offline’ planning, e.g., through a standard AI method for search such as A*, that is then followed by an execution phase, or updated ‘online’ according to perceived changes in the environment, as is the case for D* and Anytime D*, see Siegwart et al. (2011) and Russell and Norvig (1995) for a summary of approaches. Following the search/planning phase, execution typically follows *path extraction* according to the desired optimality criteria. Such a graph-based approach is here used as a baseline reference, see Section 5.1 and Appendix B for details.

An alternative approach to graph construction/search for dealing with reactive-global navigation problems is that of potential field planning (Khatib, 1986). In this case, gradient valuations that cover the entirety of the state space are composed, allowing the robot to react to unexpected obstacles or changes in the environment while still being able to navigate towards the goal by following the gradient in relation to the new state it finds itself in. Gaussier and colleagues have produced much work in the area of sequential behaviour planning in relation to navigation including adopting a potential field guided perspective. The general approach followed by this group is that of planning that back-chains potential navigable routes (sequences of spatial behavioural transitions) from a goal state to the current state. In early work, Gaussier et al. (2000) (cf. also Gaussier and Zrehen 1995, and later related work, e.g. Hirel et al. 2013) developed a neurally inspired architecture for motivated navigation. Using this approach, the robot may explore the environment and identify landmarks, corresponding to place cells (O’Keefe and Nadel, 1978), with an activation level reflecting the robot’s distance from the identified place. By propagating activity from a goal, through the network of identified landmarks, a potential field is generated, allowing the robot to navigate towards the goal by following the gradient. This approach is shown to afford performance-wise robust responding to obstacles and temporary occlusions as robot navigation is guided by the potential field (Gaussier et al., 2000).

In contrast to potential fields that implement a static attractor basin centered at the goal, the approach presented here relies on dynamic attractors forming as activity propagates from the goal to the agent’s current state. Furthermore, while potential fields use the gradient for action selection, the framework presented here relies on several discrete elementary behaviors, competing for activation. Potential fields may produce local minima in complex environments. Although not

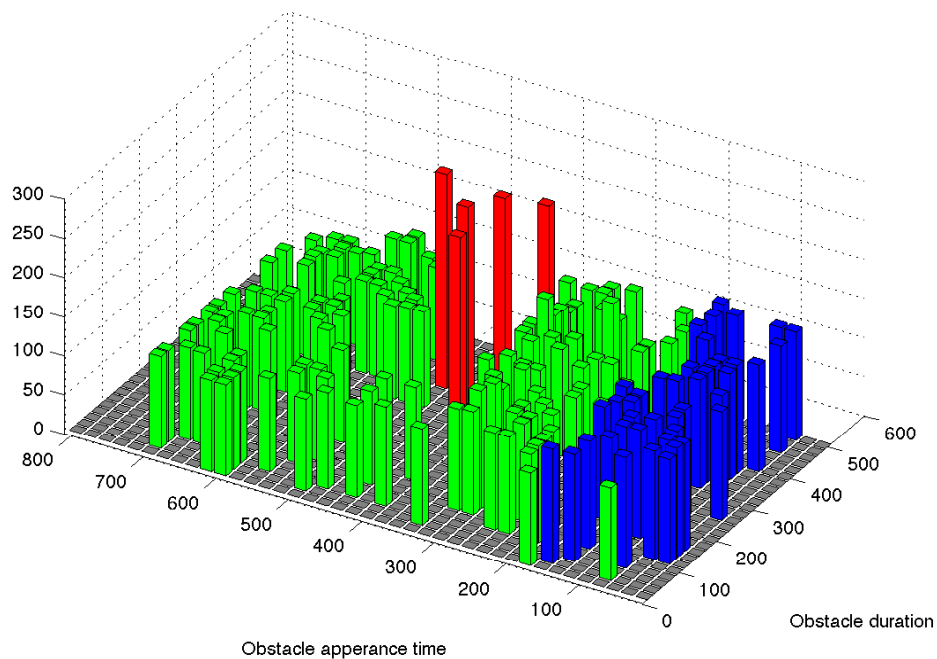


Figure 16. Path length plot over runs in Z-maze with low noise, where a partly obstructing obstacle was introduced at a random time (x-axis), for a random duration (y-axis). Bars are coloured according to selected path, as illustrated in Figure 14.

absent from local minima, the competing actions approach used here appears less prone to these problems. We aim to conduct a direct comparison with potential fields in future work.

Further alternatives to classical path planning approaches include reaction-diffusion methods (Trevai et al., 2002; Adamatzky et al., 2003; Vazquez-Otero et al., 2012) and so-called activation-diffusion (Martinet et al., 2008). In the case of reaction-diffusion, a typical approach is for the path planner algorithm to generate diffusive waves of (e.g. neurochemical) activation in a map, according to the interaction of two or more variables, as guided by mapped obstacles. Some such approaches use highly biologically focused algorithms (Adamatzky et al., 2003), others seek rather to profit from certain principles of reaction-diffusion dynamics. In the latter case, Vazquez-Otero et al. (2012) used an algorithm that firstly generated diffusive activity from a start state over a map of a maze (propagation phase) and secondly, on arriving at the goal state, activated a contraction phase (dissipation of activity), and finally, prior to action execution, a path extraction phase using a standard search algorithm. They compared their approach to classical approaches - Distance Transform (DT) and Voronoi diagrams - to assess the optimality and safety-value of the paths selected. A benefit of the approach was the creation of *smooth* paths that permit more efficient online behavior.

A general problem for reaction-diffusion mechanisms concerns finding appropriate mechanisms for permitting online reaction to changes in the environment (Adamatzky et al., 2003) and, relatedly, the dependence on explicit path extraction methods to guide online behaviour. Martinet et al. (2008) put forward an activation-diffusion method by which simulated cortical columns propagate back-chained activity from a ‘motivation’-gated column to one associated with a start state. The activation of the latter column is then forward propagated as a path signal. This activation-diffusion approach has several similarities with ours in that activation is back-chained from goal to start state and that action choices compete for activation. However, the approach presented here does not require forward propagation or the extraction of an explicit path signal.

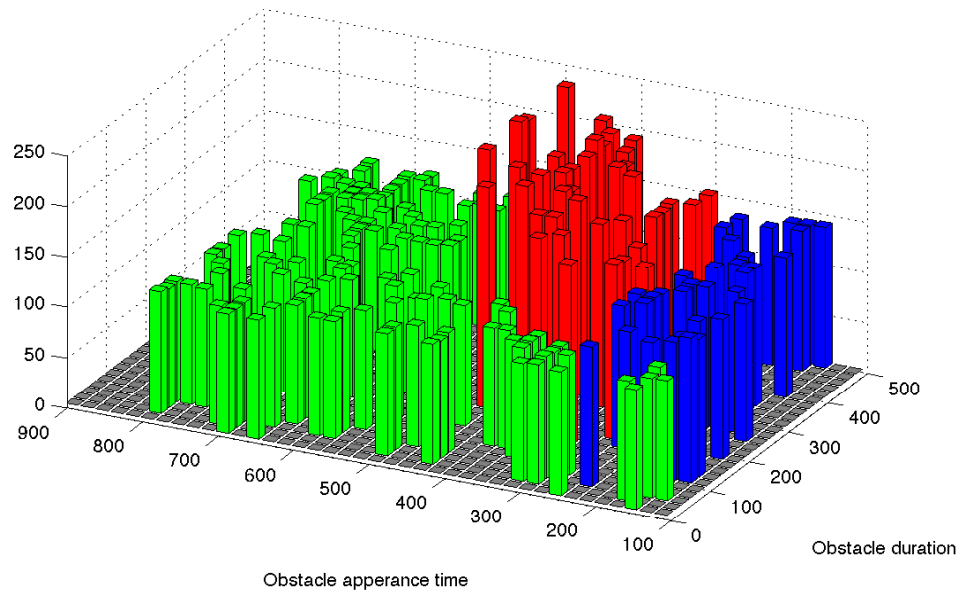


Figure 17. Path length plot over runs in the Z-maze with low noise, where a fully obstructing obstacle was introduced at a random time (x-axis), for a random duration (y-axis). Bars are coloured according to selected path, as illustrated in Figure 14.

6.2. Relation to existing planning and search methods

The properties of the neural-dynamic Elementary Behaviours, which are the building blocks of the SPA framework introduced here, are put in relation to existing planning and search algorithms:

1. *Wavefront search*: The neural dynamic nature of planning may be compared to a wavefront (i.e. multi-directional) approach. The planner is ‘systematic’ (LaValle, 2006, p. 32) - the neural dynamic nature of our planner algorithm ensures, given enough time, that all reachable states will be searched and not re-searched (it is not redundant).
2. *Shortest path search*: Initial search is Dijkstra-like - whilst neural dynamic search is carried out, multiple candidate paths are evaluated in parallel. A short path stochastically receives greater activation faster, compared to longer paths, which leads to the inhibition of competing paths.
3. *Uni-directional*: The search is uni-directional in the sense that an active region of one motivation field is partly self-sustaining and will inhibit competing transitions, i.e., re-evaluations, as long as that location is active. The self-sustainability is however dependent on neighbouring regions being active. The introduction of an obstacle, cutting of an active path, will consequently lead to a spreading deactivation of remaining parts of that route. Deactivation also releases the inhibition of competing transitions, allowing for re-evaluations of that part of the search space. This spreading activation/deactivation mechanism results in continuous adaptation to changes in the environment.
4. *Partial updates*: Search is also online and can be compared to D*. When changes in the environment occur (see Section 5.2), neural dynamic planning is updated locally (in relation to the changed aspect of the environment). This means that the robot is not required to completely replan when a path is required to be modified rather than ‘catastrophically’ altered. This is very much a feature of the SPA nature of the neural-dynamic architecture. This property shows up in the relatively short replanning times (visible as periods of low velocity in Figure 15), compared to the initial planning phase (i.e., the initial propagation of activity from the goal to the starting location of the robot).

5. *Safety*: It was found that the architecture generated navigable paths equi-distant between obstacles/walls. The path search, therefore, can be seen as an emergent Voronoi diagram. This can be put in relation to the reaction-diffusion algorithm of Vazquez-Otero et al. (2012), which produced navigable trajectories that were not purely dedicated to shortest-path exploration, but also included aspects of smoothness and a safe distance to obstacles. The approach presented here shows similar properties in that a path with longer distance to walls is selected, when that space is available. The south (blue) route depicted in Figure 14 shows a relatively long distance to all walls, compared to the north routes (green) of Case 3A where the partial obstacle only leaves a narrow corridor. Furthermore, the narrow passage leads to a reduction in velocity (Figure 15).
6. *Gradient search*: Search implicitly follows a gradient (in relation to the euclidean distance from the goal state to feasible intermediate states to the current state). However, there is no explicit gradient valuation over state space, from which potential fields can be derived. One advantage of this approach is that long paths do not suffer from low gradient valuations that are susceptible to noisy evaluations or otherwise fading memory of paths that may lead to execution problems.

6.3. Biological mechanisms for planning and navigation

Ever since the work by O'Keefe and Nadel (1978), pyramidal neurons in hippocampus with firing patterns that correlate with the physical location of the rat, so called *place cells*, have been a hot topic in cognitive neuroscience. A growing body of literature tells an increasingly complex story that involves not only place cells, but also cells in postsubiculum that are sensitive to direction (*head direction cells*) (Taube et al., 1990a,b) and cells in medial entorhinal cortex with place sensitive firing patterns arranged in hexagonal grids (*grid cells*) (Fyhn et al., 2004; Hafting et al., 2005; Sargolini et al., 2006). Hok et al. (2005) have also provided evidence for cells in medial prefrontal cortex (mPFC) that reflects the motivational salience of places (*goal cells*).

Experimental research on rodents has inspired a large body of computational investigations, e.g. Burgess et al. (1994) and Redish and Touretzky (1997). These are also models, inspired by research on place cells, directed towards robotic applications targeting the problem of simultaneous localisation and mapping (SLAM). A biologically inspired model, RatSLAM (Milford and Wyeth, 2008), has shown impressive results in mapping 66 km of urban roads using only a single camera. An algorithm like RatSLAM could potentially be linked with the architecture presented here via the *Place sense* and *Head direction* fields (Fig. 5).

While place cells as a mechanism for self localization and mapping appear to be relatively well understood, their role in path planning and goal pursuit is still puzzling (Jeffery et al., 2003; Poucet et al., 2004). Computational models of navigation based on place cells typically formulate the problem using reinforcement learning, using both classical artificial neural networks (Kulvicius et al., 2008) and spiking networks (Strösslín et al., 2005). Available paths have been modeled as transitions from one place (represented by a place cell's firing field) to another, and coded as *transition cells*, corresponding to CA3 pyramidal cells (Banquet et al., 2002; Gaussier et al., 2002). The present work takes inspiration from this view of place cells, reflected in our definition of elementary behaviours as leading to a transition in a particular direction (see Section 3 for details).

The work presented here can also be put in relation to the parallel view of perception and action proposed by Cisek and Kalaska (2010). Cisek and Kalaska summarize a large body of neurophysiological evidence speaking for a view where action selection and parameterization are processed in parallel to perception. In this view, neural activity related to different decisions (action responses) builds up as a function of motivational value and perceptual evidence, speaking for or against different responses. While the architecture presented here is not aimed at closely resembling any biological mechanism, it constitutes one example of a computational model that captures many of these properties. Not only selection of short

term action responses but also planning is in this view seen as preparation of competing actions, continuously influenced by perceptual evidence and motivational value.

6.4. Effects of noise

In the gradient-based approach to planning sequential behaviour used by Gaussier and colleagues it has been suggested that long paths may be difficult to plan – “if we want our animat to learn paths that need more than several tens of subgoals then the slope of the gradient [from goal representation to representation of current state] will be very low ... and will be very difficult to use ... There is a need to be able to structure the plans ...” (Gaussier et al., 2000, p. 86). In general, it has been suggested (Koren and Borenstein, 1991) that potential field approaches to robot navigation tasks suffer many limitations. For example, robots may a) get stuck in local minima, b) get stuck between non-passable objects/obstacles, c) oscillate/dither in the presence of obstacles or narrow passages.

Some of these limitations appear to be, at least partly, present in the architecture presented here. The robot does, on a few occasions, get stuck in local minima, (see Section 5.1 for details). One such example is displayed in Figure 10, the red colored run with low noise. In this case, there are two competing attractors, one from the south transition network and one from the east. Since the distance to the goal happens to be almost exactly the same at the point of the junction, the two attractors are equally strong and, in combination with the inhibitory inflow from the obstacle avoidance, a local minimum appears.

One common way to handle local minima problems in potential fields is by introducing noise. As has been demonstrated in Section 5, the architecture presented here appears very robust to noise. Even though the high noise level did not lead to 100% successful runs in all cases, it did reduce the number of unsuccessful runs. The fact that no significant effect of noise was found in Cases 1 and 2 indicates that noise can be introduced with a very small cost in terms of navigation performance. These results should however be taken as preliminary, a deeper analysis of the system’s behavior in response to noise is necessary to provide a full understanding of this potentially beneficial property.

6.5. Learning and development

While learning is not studied experimentally here, it is still an important aspect of the proposed architecture. The architecture has many parameters and it is therefore desirable to show how these could emerge from learning.

In the present context, learning can be considered on at least three levels of abstraction: 1) learning individual EBs, 2) learning relations between EBs, and 3) learning the constraints imposed by a particular environment.

The first problem may be addressed by combining reinforcement learning with the DNF framework. Based on the same principles as the used in the architecture presented here, Kazerounian et al. (2012) demonstrated one such example. Outside the DFT framework, we have also studied learning of EBs in a navigational context. Billing et al. (2011, 2015) applied the TD-algorithm *Predictive Sequence Learning (PSL)* to sensor and motor data from a Robosoft Kompai robot, allowing the robot to learn and generalize goal directed actions demonstrated via teleoperation.

We have also studied the second form of learning as the learning of connection weights between EBs (Sandamirskaya and Schöner, 2010; Luciw et al., 2013). This should be understood as learning how behavioral consequences relate to each other and are typically environment independent. In the present work, we provide suitable connection weights between EBs, as described in Section 3.

Learning behavioral constraints for a particular environment can be seen as a SLAM problem. As discussed in Section 6.3, Milford and Wyeth (2008) propose an approach with several connections to the present work. The presented technique, called RatSLAM, implements *Continuous Attractor Networks (CAN)* with similar dynamics to the DNFs used in the present architecture. In future work, we aim to combine these pieces of work in order to also learn the contextual input, i.e., the environment map, during simultaneous planning action.

7. Conclusion

We have presented an approach to planning that exploits *Dynamic Field Theory (DFT)* based attractor networks to produce robust sequencing of elementary behaviors with an online updating of the planned path. We have implemented the developed approach in an exemplary path planning and navigation scenario. The developed architecture consists of a tuple of *Dynamic Neural Fields (DNFs)*, in which activity spreads to propagate the motivational input (goal information), while prior knowledge of the environment is represented with an inhibitory flow to the motivation DNFs, and between the precondition and intention DNFs. The architecture receives direct inputs from the sensors and controls actuators of the robot, producing a system that performs neural-dynamic search in order to discover available paths towards the goal, whilst simultaneously permitting behaviour. We refer to this property as *Simultaneous Planning and Action (SPA)*, effectively a form of online search as actions and changes in the environment perturb existing path plans.

In summary, the evaluation of our architecture in a simulated robotic experiments demonstrated the following features:

- *Scalability*: The proposed architecture scales well with the size or complexity of the environment. The activation spreads with a constant speed among DNFs and eventually reaches the current location of the agent, creating a landscape of attractors that leads towards the goal.
- *Noise Robustness*: The proposed SPA architecture appears robust to noise. Increased noise does not have any significant effect on navigation performance and does not appear to affect planning.
- *Performance robustness*: Thanks to the stability of attractor dynamics, transient changes and blockages of the environment do not result in replanning. The selected path will however be abandoned if an obstacle is blocking the way for a prolonged amount of time. Depending on the duration and extent of the blockage, an alternative route will be selected.

Of critical importance to the performance of the SPA architecture is its robust online path planning behavior. Noise and performance robustness are continuous in that minor perturbations (including partial and/or transient obstructions) do not deter the agent from following its goal path. Major perturbations, however, lead to path plan updates and the following of an alternative route to the goal.

We demonstrate that the proposed architecture can produce path planning behavior comparable to traditional methods based on search (see specifically Figures 8 and 9). The number of successful runs in the larger environment is notably lower for the ‘Follow the carrot’, compared to the proposed SPA architecture. This is a result of the ‘Follow the carrot’ algorithm cutting corners and therefore getting stuck at tight turns. While this is a common problem in classical approaches, the competition between actions that occur in SPA does not appear to cut corners. In addition, the high success rate for the proposed architecture partially may be attributed to tight integration of input from the proximity sensors, which biases action selection, delaying a turn-action if a wall is blocking the path. It should, however, be noted that there are other classical path tracking methods less prone to cut corners, and the problem can also be reduced using a reactive obstacle avoidance layer. Hence, these results should not be understood as a limitation of traditional search methods in general.

The approach to simultaneous planning and action presented here shows robustness on multiple levels, from emergent decision making and high noise tolerance, to stable interactions with the world, without the need to separate the system into a deliberative and a reactive layers. While this approach still needs more analysis before it can be considered as a mature alternative to traditional planning, the present work should be seen as a proof of concept for dynamical approaches, and specifically DFT, applied as a method for planning and search.

Acknowledgment

This work was funded through the 7th framework of the EU in grant #270247 (*Neural Dynamics*). We would also like thank Sebastian Schneegans, Institute for Neural Computation, Ruhr-Universität Bochum, for providing the software framework COSIVINA (Schneegans, 2015) and invaluable help with technical support.

Appendix A. Mathematical formulation

Dynamic neural fields

The activation level u of the neural fields constituting the M, P, I, and C components (Section 3) evolves according to Equations 4, 5, 6, and 7, respectively:

$$\begin{aligned} \tau \dot{u}_M(x, t) = & -u_M(x, y, t) + h + \\ & \int f(u_M(x', y', t)) \omega(x - x', y - y') dx' dy' + \\ & S_2(x, y, t) - S_\gamma(x, y, t), \end{aligned} \quad (4)$$

$$\begin{aligned} \tau \dot{u}_P(x, t) = & -u_P(x, y, t) + h + \\ & \int f(u_P(x', y', t)) \omega(x - x', y - y') dx' dy' + \\ & S_1(x, y, t) - S_7(x, y, t), \end{aligned} \quad (5)$$

$$\begin{aligned} \tau \dot{u}_I(x, t) = & -u_I(x, y, t) + h + \\ & \int f(u_I(x', y', t)) \omega(x - x', y - y') dx' dy' + \\ & S_3(x, y, t) - S_4(x, y, t) - S_6(x, y, t), \end{aligned} \quad (6)$$

$$\begin{aligned} \tau \dot{u}_C(x, t) = & -u_C(x, y, t) + h + \\ & \int f(u_C(x', y', t)) \omega(x - x', y - y') dx' dy' + \\ & S_5(x, y, t) + S_8(x, y, t). \end{aligned} \quad (7)$$

ω is a two-dimensional version of the standard gaussian kernel given by Equation 2. Table 1 specifies connections S_1 to S_8 , also visible in Figures 2 and 5.

Action fields

The action field of each EB (Fig. 5) is defined over an angular dimension, from $-\pi$ to π , with a circular mapping, such that $u(x, t) = u(x + 2\pi, t)$, for all x . The connections (12 and 13) from the action field to the left and right wheels cause the robot to turn to the left (μ_{lw}) and to the right (μ_{rw}) in response to a peak on the negative and positive sides of the action field, respectively. The wheel speeds are given by:

$$\mu_{lw} = \vartheta_{max}(x, \alpha) + \vartheta_{min}(x, \alpha), \quad (8)$$

$$\mu_{rw} = \vartheta_{max}(x, -\alpha) + \vartheta_{min}(x, -\alpha), \quad (9)$$

$$\vartheta_{max}(x, \alpha) = \max_{x=-\pi}^{\pi} f(u_a(x, t)) \sin(x + \alpha) v, \quad (10)$$

$$\vartheta_{min}(x, \alpha) = \min_{x=-\pi}^{\pi} f(u_a(x, t)) \sin(x + \alpha) v, \quad (11)$$

where v and α are constants controlling the speed and turning rate of the robot, respectively. f is the field output function given by Equation 3 and u_a represents the activation level of the action field, given by:

$$\begin{aligned} \tau \dot{u}_a(x, t) = & -u_a(x, t) + \int f(u(x', t)) \omega(x - x') dx' \\ & + h + S_I(t) + S_H(x, t) - S_O(x, t), \end{aligned} \quad (12)$$

where

Table 1. Connectivity functions

$$\begin{aligned}
S_1(x, y, t) &= c_1 \int f(u_M(x', y', t)) g(x - x' + \delta x, y - y' + \delta y) dx dy, \\
S_2(x, y, t) &= c_2 \int f(u_P(x', y', t)) g(x - x', y - y') dx dy, \\
S_3(x, y, t) &= c_3 \int f(u_M(x', y', t)) g(x - x', y - y') dx dy, \\
S_4(x, y, t) &= c_4 \int f(u_P(x', y', t)) g(x - x' + \delta x, y - y' + \delta y) dx dy, \\
S_5(x, y, t) &= c_5 \int f(u_I(x', y', t)) g(x - x', y - y') dx dy, \\
S_6(x, y, t) &= c_6 \int f(u_C(x', y', t)) g(x - x', y - y') dx dy, \\
S_7(x, y, t) &= c_7 \int f(u_C(x', y', t)) g(x - x' + \delta x, y - y' + \delta y) dx dy, \\
S_8(x, y, t) &= c_8 g(x - x_{robot}(t), y - y_{robot}(t)), \\
g(\Delta x, \Delta y) &= \exp \left[-\frac{\Delta x^2 + \Delta y^2}{2\sigma^2} \right].
\end{aligned}$$

$$S_I(t) = \max_x [f(u_I(x, t))], \quad (13)$$

$$S_H(x, t) = f(u_H(x' + \theta, t)) \omega(x - x') dx', \quad (14)$$

$$S_O(x, t) = f(u_O(x', t)) \omega(x - x') dx'. \quad (15)$$

S_I , S_H , and S_O are the stimuli received from the corresponding I-field, the head direction field, and the obstacle field, respectively. u_I , u_H , and u_O are activation functions for the respective fields, given by Equation 1. θ is a directional constant, shifting the input in relation to action field direction. The interaction kernel ω is given by Equation 2.

Proximity sensors

The activation level of the proximity sensor field (Figure 5) is given by:

$$\begin{aligned}
\tau \dot{u}_O(x, t) &= \\
&= -u(x, t) + h + \int f(u(x', t)) \omega(x - x') dx' + \sum_r S_r(x, t).
\end{aligned} \quad (16)$$

S_r is the stimuli from each proximity sensor r , given by:

$$S_r(x, t) = \log[1 + p_r(t)/50] c_{prox} \exp \left[-\frac{(x - \check{x}_r)^2}{2\sigma_{prox}^2} \right], \quad (17)$$

where $p_r(t)$ is the sensor value of proximity sensor r at time t . c_{prox} and σ_{prox} are the constants controlling the strength and standard deviation of S_r , respectively. See Table 2 for details.

Table 2. List of variables and constants. See also Figure 5 for details on field connectivity.

Symbol	Description	Value			
u	Activation level of a neural field				
x	Position in a neural field				
c_{1-8}	Amplitudes for connections (Table 1). $c_1:12, c_{2a}:5, c_{2b}:-4.5, c_{2c}:4.5, c_{2d}:-10, c_3:10, c_4:-10, c_5:3, c_6:-5, c_7:-15, c_8:8$.				
t	Time specified in simulation steps (64 ms/step)				
τ	Time constant controlling field dynamics				5
h	Field resting level				-5
σ	Variance of ω , controlling the width of the excitatory part of the interaction kernel				2
v	Field output to speed conversion (tics/s)				500
α	Constant controlling angular speed				$\pi/3$
c_{prox}	Strength of e-puck proximity sensor stimuli				5
σ_{prox}	Standard deviation of the e-puck proximity sensor stimuli				3
\tilde{x}_r	Field position of stimuli to the proximity sensor field (Eq. 16 and 17): $\tilde{x}_0 = 0.1\pi, \tilde{x}_1 = 0.3\pi, \tilde{x}_2 = 0.8\pi, \tilde{x}_3 = 1.0\pi, \tilde{x}_4 = -1.0\pi, \tilde{x}_5 = -0.8\pi, \tilde{x}_6 = -0.3\pi, \tilde{x}_7 = -0.1\pi$. The angular field position reflects the physical placement of sensors on the e-puck robot.				
Constants dependent on field type: motivation (M), precondition (P), intention (I) and condition of satisfaction (C)					
c_{exc}	Amplitude of the excitatory component of ω	7	4	9	2
σ_{exc}	Standard deviation of the excitatory component of ω	2	2	2	2
c_{inh}	Amplitude of the inhibitory component of ω	2	2	0	2
σ_{inh}	Standard deviation of the inhibitory component of ω	4	4	4	4
c_{global}	Global inhibition	0	0	0	0
Constants dependent on transition direction, north (N), south (S), west (W) or east (E)					
δx		N	S	W	E
δy	Connection shift (corresponding to the translation of the field)	0	0	5	-5
		5	-5	0	0

Appendix B. Follow the carrot

Follow the carrot (e.g. Barton, 2001) is a simple method for path tracking. As a baseline implementation, Follow the carrot was used as a comparison in the evaluation of the the proposed architecture.

A steering angle ϕ_{ftc} is given by $\phi_{ftc} = k_{ftc}(\phi_c - \phi_h)$, where ϕ_c is the angle between robot and selected carrot point, and ϕ_h is the heading angle.

The carrot point p_c is defined as:

$$p_c = \min_{p \in P} [d(p, p_{look})] \quad (18)$$

where P is the set of points on the path, and p_{look} is the lookahead point at distance $k_{look} = 1cm$, directly in front of the robot. $d(p, p_{look})$ denotes the Euclidean distance between p and p_{look} .

The path P is given by A* over an 8 neighbor grid with a cell size of 1 cm and a minimum obstacle distance of 4 cm. This path calculation is also used as a reference *optimal path*, Section 5.

Follow the

References

- Adamatzky, A., De Lacy Costello, B., Melhuish, C., and Ratzcliffe, N. (2003). Experimental Reaction-Diffusion Chemical Processors for Robot Path Planning. *Journal of Intelligent and Robotic Systems*, 37:233–249.
- Amari, S. (1977). Dynamics of Pattern Formation in Lateral-Inhibition Type Neural Fields. *Biological Cybernetics*, 27(2):77–87.
- Arbib, M. A. (1985). Schemas for the temporal control of behavior. *Hum Neurobiol*, 4(2):63–72.
- Arkin, R. C. (1998). *Behaviour-Based Robotics*. MIT Press.
- Banquet, J. P., Gaussier, P., Quoy, M., Revel, A., and Burnod, Y. (2002). Cortico-hippocampal maps and navigation strategies in robots and rodents. In *ICSAB Proceedings of the seventh international conference on simulation of adaptive behavior on From animals to animats*, pages 141–150, MA, USA. MIT Press Cambridge.
- Barton, M. J. (2001). *Controller Development and Implementation for Path Planning and Following in an Autonomous Urban Vehicle*. Phd thesis, School of Aerospace, Mechanical and Mechatronic Engineering, The University of Sydney, Australia.
- Bastian, A., Schöner, G., and Riehle, A. (2003). Preshaping and continuous evolution of motor cortical representations during movement preparation. *European Journal of Neuroscience*, 18(7):2047–2058.
- Billing, E. A. and Hellström, T. (2010). A Formalism for Learning from Demonstration. *Paladyn: Journal of Behavioral Robotics*, 1(1):1–13.
- Billing, E. A., Hellström, T., and Janlert, L. E. (2011). Simultaneous Control and Recognition of Demonstrated Behavior. Technical report, Department of Computing Science, Umeå, Sweden.
- Billing, E. A., Hellström, T., and Janlert, L. E. (2015). Simultaneous Recognition and Reproduction of Demonstrated Behavior. *Biologically Inspired Cognitive Architectures (in press)*.
- Brooks, R. A. (1991). New approaches to robotics. *Science*, 253:1227–1232.
- Burgess, N., Recce, M., and O’Keefe, J. (1994). A model of hippocampal function. *Neural networks*, 7:1065–1081.
- Cisek, P. and Kalaska, J. F. (2010). Neural mechanisms for interacting with a world full of action choices. *Annual review of neuroscience*, 33(March):269–98.
- Cyberbotics (2015). Webots.
- Ermentrout, B. (1998). Neural networks as spatio-temporal pattern-forming systems. *Reports on Progress in Physics*, 61(4):353–430.
- Fyhn, M., Molden, S., Witter, M. P., Moser, E. I., and Moser, M.-B. (2004). Spatial representation in the entorhinal cortex. *Science (New York, N.Y.)*, 305(5688):1258–64.
- Gaussier, P., Lepître, S., Quoi, M., Revel, A., and Joulain, C. (2000). Experiments and Models about Cognitive Map Learning and Motivated Navigation. In Demiris, J. and Birk, A., editors, *Interdisciplinary Approaches to Robot Learning*, pages 53–94.
- Gaussier, P., Revel, A., Banquet, J. P., and Babeau, V. (2002). From view cells and place cells to cognitive map learning: processing stages of the hippocampal system. *Biological cybernetics*, 86(1):15–28.
- Gaussier, P. and Zrehen, S. (1995). Perac: A neural architecture to control artificial animals. *Robotics and Autonomous Systems*, 16:291–320.
- Hafting, T., Fyhn, M., Molden, S., Moser, M.-B., and Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature*, 436(7052):801–806.
- Hirel, J., Gaussier, P., Quoy, M., Banquet, J. P., Save, E., and Poucet, B. (2013). The hippocampo-cortical loop: spatio-temporal learning and goal-oriented planning in navigation. *Neural networks : the official journal of the International Neural Network Society*, 43:8–21.
- Hok, V., Save, E., Lenck-Santini, P. P., and Poucet, B. (2005). Coding for spatial goals in the prelimbic/infralimbic area of the rat frontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 102(12):4602–7.
- Jeffery, K. J., Gilbert, A., Burton, S., and Strudwick, A. (2003). Preserved performance in a hippocampal-dependent spatial task despite complete place cell remapping. *Hippocampus*, 13(2):175–89.

- Johnson, J. S., Spencer, J. P., and Schöner, G. (2008). Moving to higher ground: The dynamic field theory and the dynamics of visual cognition. *New Ideas in Psychology*, 26:227–251.
- Johnson, J. S., Spencer, J. P., and Schöner, G. (2009). A layered neural architecture for the consolidation, maintenance, and updating of representations in visual working memory. *Brain research*, 1299:17–32.
- Kazerounian, S., Luciw, M., Richter, M., and Sandamirskaya, Y. (2012). Autonomous reinforcement of behavioral sequences in neural dynamics. In *Proceedings of the Joint IEEE International Conference on Development and Learning & Epigenetic Robotics (ICDL-EPIROB)*.
- Khatib, O. (1986). Real-time obstacle avoidance for manipulators and mobile robots. *The international journal of robotics research*, 5:90–98.
- Koren, Y. and Borenstein, J. (1991). Potential Field Methods and Their Inherent Limitations for Mobile Robot Navigation. In *IEEE International Conference on Robotics and Automation*, pages 1398–1404, Sacramento, California.
- Kulvicius, T., Tamosiunaite, M., Ainge, J., Dudchenko, P., and Wörgötter, F. (2008). Odor supported place cell model and goal navigation in rodents. *Journal of computational neuroscience*, 25(3):481–500.
- LaValle, S. M. (2006). *Planning Algorithms*. Cambridge University Press, Cambridge, U.K.
- Luciw, M., Kazerounian, S., Lakhmann, K., Richter, M., and Sandamirskaya, Y. (2013). Learning the perceptual conditions of satisfaction of elementary behaviors. In *Robotics: Science and Systems (RSS), Workshop "Active Learning in Robotics: Exploration, Curiosity, and Interaction"*.
- Martinet, L.-E., Passot, J.-B., Fouque, B., Meyer, J.-A., and Arleo, A. (2008). Map-Based Spatial Navigation: A Cortical Column Model for Action Planning. In et al. Freksa, C., editor, *Spatial Cognition VI. Learning, Reasoning, and Talking about Space.*, number 1, pages 39–55. Springer Berlin Heidelberg.
- Matarić, M. J. (1997). Behavior-Based Control: Examples from Navigation, Learning, and Group Behavior. *Journal of Experimental and Theoretical Artificial Intelligence*, 9(2-3):323–336.
- Matarić, M. J. (2000). Sensory-motor primitives as a basis for imitation: linking perception to action and biology to robotics. In *Imitation in animals and artifacts*, pages 391–422. MIT Press.
- Milford, M. and Wyeth, G. (2008). Mapping a Suburb With a Single Camera Using a Biologically Inspired SLAM System. *IEEE Transactions on Robotics*, 24(5):1038–1053.
- Mondada, F., Bonani, M., Raemy, X., Pugh, J., Cianci, C., Klapotcz, A., Magnenat, S., Zufferey, J.-C., Floreano, D., and Martinoli, A. (2009). The e-puck, a Robot Designed for Education in Engineering. In *Proceedings of the 9th Conference on Autonomous Robot Systems and Competitions*, pages 59–65.
- Nicolescu, M. (2003). *A Framework for Learning from Demonstration, Generalization and Practice in Human-Robot Domains*. PhD thesis, University of Southern California.
- O’Keefe, J. and Nadel, L. (1978). *The Hippocampus as a Cognitive Map*. Oxford University Press, Oxford.
- Poucet, B., Lenck-Santini, P. P., and Hok, V. (2004). Spatial navigation and hippocampal place cell firing - the problem of goal encoding. *Reviews in the Neurosciences*, 15:89–107.
- Redish, a. D. and Touretzky, D. S. (1997). Cognitive maps beyond the hippocampus. *Hippocampus*, 7(1):15–35.
- Richter, M., Sandamirskaya, Y., and Schöner, G. (2012). A robotic architecture for action selection and behavioral organization inspired by human cognition. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS*.
- Russell, S. and Norvig, P. (1995). *Artificial Intelligence: A Modern Approach*. Prentice Hall, NJ, 1 edition.
- Sandamirskaya, Y., Richter, M., and Schöner, G. (2011). A neural-dynamic architecture for behavioral organization of an embodied agent. *EEE International Conference on Development and Learning (ICDL)*, 2:1–7.
- Sandamirskaya, Y. and Schöner, G. (2010). An embodied account of serial order: How instabilities drive sequence generation. *Neural Networks: The Official Journal of the International Neural Network Society*, 23(10):1164–1179.
- Sandamirskaya, Y., Zibner, S., Schneegans, S., and Schöner, G. (2013). Using dynamic field theory to extend the embodiment stance toward higher cognition. *New Ideas in Psychology. Special Issue on Adaptive Behavior*.

- Sargolini, F., Fyhn, M., Hafting, T., McNaughton, B. L., Witter, M. P., Moser, M.-B., and Moser, E. I. (2006). Conjunctive representation of position, direction, and velocity in entorhinal cortex. *Science (New York, N.Y.)*, 312(5774):758–62.
- Schneegans, S. (2015). COSIVINA (Compose, Simulate, and Visualize Neurodynamic Architectures). <http://roboticsschool.ini.rub.de/software.php>.
- Schöner, G. (2008). Dynamical systems approaches to cognition. In *Cambridge handbook of computational cognitive modeling*, pages 101–126. Cambridge University Press, Cambridge, UK.
- Schöner, G., Kopecz, K., and Erlhagen, W. (1997). The dynamic neural field theory of motor programming: Arm and eye movements. In Morasso, P. G. and Sanguineti, V., editors, *Self-Organization, Computational Maps and Motor Control, Psychology Series, Vol. 119*, pages 271–310. Elsevier-North Holland.
- Siegwart, R., Nourbakhsh, I., and Scaramuzza, D. (2011). *Introduction to Autonomous Mobile Robots*. The MIT Press, 2 edition.
- Spencer, J. P., Perone, S., and Johnson, J. S. (2009). Dynamic Field Theory and Embodied Cognitive Dynamics. In Spencer, J., Thomas, M., and McClelland, J., editors, *Toward a Unified Theory of Development: Connectionism and Dynamic Systems Theory Re-Considered*, pages 86–118. Oxford University Press., New York.
- Spencer, J. P. and Schöner, G. (2003). Bridging the representational gap in the dynamical systems approach to development. *Developmental Science*, 6:392–412.
- Strösslín, T., Sheynikhovich, D., Chavarriaga, R., and Gerstner, W. (2005). Robust self-localisation and navigation based on hippocampal place cells. *Neural networks : the official journal of the International Neural Network Society*, 18(9):1125–1140.
- Tani, J., Ito, M., and Sugita, Y. (2004). Self-Organization of Distributedly Represented Multiple Behavior Schemata in a Mirror System : Reviews of Robot Experiments Using RNNPB. *Neural Networks*, 17:1273–1289.
- Taube, J. S., Muller, R. U., and Ranck, J. B. (1990a). Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 10(2):420–35.
- Taube, J. S., Muller, R. U., and Ranck, J. B. (1990b). Head-direction cells recorded from the postsubiculum in freely moving rats. II. Effects of environmental manipulations. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 10(2):436–47.
- Tolman, E. C. (1948). Cognitive maps in rats and men. *The Psychological Review*, 55(4).
- Trevai, C., Fukazawa, Y., Yuasa, H., Ota, J., Arai, T., and Asama, H. (2002). Cooperative exploration path planning for mobile robots by reaction-diffusion equation on graph. In *IEEE International Conference on Industrial Technology (ICIT)*, volume 2, pages 1266–1271.
- Vazquez-Otero, A., Faigl, J., and Muñozuri, A. P. (2012). Path planning based on reaction-diffusion process. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 896–901, Vilamoura, Algarve, Portugal.
- Wilson, H. R. and Cowan, J. D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik*, 13(2):55–80.
- Zibner, S. K. U., Faubel, C., Iossifidis, I., and Schöner, G. (2011). Dynamic neural fields as building blocks for a cortex-inspired architecture of robotic scene representation. *IEEE Transactions on Autonomous Mental Development*, 3(1):74–91.